Journal of Theoretical and Applied Information Technology
15th August 2021. Vol.99. No 15
© 2021 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

DEVELOPMENT OF INTELLIGENT INFORMATION SYSTEM OF ASSESSING THE NEGATIVE IMPACT OF INDUSTRIAL EMISSIONS ON THE PUBLIC HEALTH

[1]OLGA SHVETS, [2]ALINA BUGUBAYEVA, [3]SAULE RAKHMETULLINA, [4]WALDEMAR WOJCIK

[1]Associate Professor, East-Kazakhstan Technical University, Department of Information technology and intelligent system, Kazakhstan

[2]PhD Student, East-Kazakhstan Technical University, Department of Information technology and intelligent system, Kazakhstan

[3]First Vice-Rector, East-Kazakhstan Technical University, University administration, Kazakhstan

[4]Professor, Lublin University of Technology (Politechnika Lubelska), University administration, Poland

E-mail: [1]olga.shvets75@gmail.com, [2]alina_bugubayeva@mail.ru, [3]SRakhmetullina@ektu.kz, [4]waldemar.wojcik@pollub.pl

ABSTRACT

The following investigation is based upon the treated patients' data of the East-Kazakhstan region and information about the level of emissions in Ust-Kamenogorsk for the same period. The analysis resulted in correlation between specific diseases and industrial emissions. The study of the influence of environmental factors on the health of the population and the risk of disease is an urgent and demanded task. The idea of the work: to identify and quantify the relationship between atmospheric pollution by emissions of harmful substances from industrial enterprises with the damage caused to the health of the population. The purpose of the work: to develop, on the basis of theoretical research and processing of statistical materials, a methodology for predicting the negative impact of harmful emissions on the health of the population. The research was constructed with neuro-net modelling technology. "Intelligent information system of Assessing the negative impact of industrial emissions on human health (on the example of Ust-Kamenogorsk data)" based on neural network technology was created to automate tasks. The information system was developed using the C # programming environment. A certificate on data registration in the state register of copyrighted objects of the Republic of Kazakhstan No. 16777, from April 20, 2021 (see Figure 10) was received.

Keywords: *Neuro-net, Modelling, Ecological Monitoring, Health, Ecological Predicting*

1. INTRODUCTION

Any ecosystem consists of a great number of dynamic processes, interconnected components of a biological, climatic, chemical, geological, physical or social nature [1]. When analyzing the state of ecosystems or problems dealing with the systems, processing the data obtained from various sources, different formats and dimensions can be difficult. Currently, the main adverse effects on human health are admitted by many international scientific organizations and most scholars dealing with the problem of low-dose biological effects [1]:

- ✓ increase in the frequency of malignant neoplasms of certain organs (or tissues);
- ✓ increase in the frequency of some hereditary diseases in offspring.

Both types of effects are stochastic (probabilistic) in nature, while the effects are very small, so they cannot be measured directly (for example, in an experiment), and the well-known dose-response or dose-effect is used to assess the effects of small doses. When used in the range of high to medium doses the ratio is extrapolated with certain hypotheses and models for the low-dose range.

This issue has been studied seriously in US medical institutions, which conducted research over many years in several hundreds of cities, separately for each industrial pollutant and for each physical effect. For example, they studied the effect of changes in the surface concentration of dust particles with a size of up to 10 μm PM10 on sudden mortality. Fine PM10 dust remains in the lungs of a person forever, in contrast to coarse dust, which is retained in the respiratory tract and excreted from the body. For example, dust particles, along with chemical compounds that pollute the air, play an important role in the occurrence of lung diseases. Lung diseases decrease as long as air pollution with the same chemical pollution decreases. [2, 3].

Achievement of such an important monitoring goal as the formation of logical conclusions and the issuance of recommendations suitable for supporting decision-making on the control of the investigated object in fuzzy conditions or recommendations in optimization or stabilization decisions requires some specific intelligent mathematical models, methods and algorithms that have not been thoroughly developed yet .

Currently, there are many methods and works on the analysis of the influence of harmful substances on the diseases, however, all works are based on the study of the influence of specific chemical (organic and inorganic) substances on the human body – on its individual organs, the immune system and on the dynamics of changes in the structure of cells depending on the dose of certain harmful substances. At the same time, despite the serious achievements obtained by calculating the concentrations of harmful substances in the urban sector and the region, the degree of their influence on the occurrence of diseases in real time has been poorly studied and has not been published in the literature.

The research aims at fully monitoring an object state which requires developing models, methods and numerical methods and state algorithms based on measuring significant object parameters including incomplete, imperfect and careless data to realize them in a decision support system.

These models can be realized through artificial intelligence systems such as methods based on fuzzy logic, neural networks, genetic algorithms and their combinations. Neural network model can be instructed by measuring data and making it possible to summarize incomplete and noisy data. Fuzzy models are able to process high-quality linguistic information, and neural networks have both of these advantages. Lately, due to complex and nonlinear processes, artificial neural networks attract the attention of ecologists and scientists dealing with scientific prognostication and computing environmental data. Since these networks tend to be used for ecological modeling, approximation by a function, forecasting, data qualification and data mining, we use this mechanism for intellectual analysis.

2. ENVIRONMENTAL IMPACT ON PEOPLE'S HEALTH AND DATA MINING

Such ecological monitoring tasks as ecological data capturing with its further computing and interpretation, ratings control, are being successfully solved. Technology of gathering and storing of information are considered by national and foreign authors as J.Park, S. McCay, V.V. Smorodin, E.V. Volkova [4, 5]. For example, statistical treatment of ecological data is considered in classic paper by M.E. Berlyand, E.A. Vukolov, Yang, Y., Sun, H., Xue, J and others. [6-8]. Ecologic data analysis is carried out with separate, disconnected methods or is limited by comparing ecological dimension measurement with marginal standards.

The questions of danger identification and risk assessment procedures in normal and emergency functioning at hazardous factories are the matter of great importance in foreign and domestic literature. Risk management issues in an industrial region are often in a conceptual stage and lack development of alternative methods and approaches in an interdisciplinary level [9, 10]. For example, the paper presents a detailed risk analysis procedure for the region [10].

GIS is able to process and analyze map material in the form of e-map of a territory and information on characteristics of the natural environment and human intervention objects. It is one of modern methods of consolidation of different-type data and integral presentation of findings. The works of the following domestic authors: G.M. Mutanov, N.M. Temirbekov, M.N. Madiyarov, F.N. Abdoldina, O.Y. Shvets, M.T. Ipalakova, U.M. Sultangazin, B.T. Zhumagulov – consider the tasks of formalization of different-type map and attributive data, their processing, further analysis and visual representation with GIS-based

Journal of Theoretical and Applied Information Technology
15th August 2021. Vol.99. No 15
© 2021 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

technologies [11-17]. Y.I Shokin, O.L. Zhizhimov, K.T. Protasov, L.F. Nozhenkova and some others are considered to be distinguished representatives from Siberian Department of RAS dealing with investigations in this field [18-20].

The question of industrial pollution impact on the environment and in particular on population health is of great practical interest for leading scientists not only from the area of geo-ecology, environmental science and medicine, but also information technology. The first papers about using this technology appeared in the USA in the 80-s of the XX century. A classic investigation is based on risk assessment of different harmful and hazardous chemical substances for human health. Since then many methods and models to define risk types and reasons for such assessment necessity have been studied and devised. Risk analysis methodology developed in the Security Center, USA (https://securitycenterusa.com/), could be taken as an example.

When developing an information system great volumes of literature [21] as well as present IT standards have been analyzed [22]. Assessment of the dependence "dose-effect" quantitative relationship between substance dose and effect on health.

Dependence assessment "dose-effect" reflects quantitative relationship between the level of influence and resulting harmful effect on level of health (properly response or reaction) [9, 23, 24].

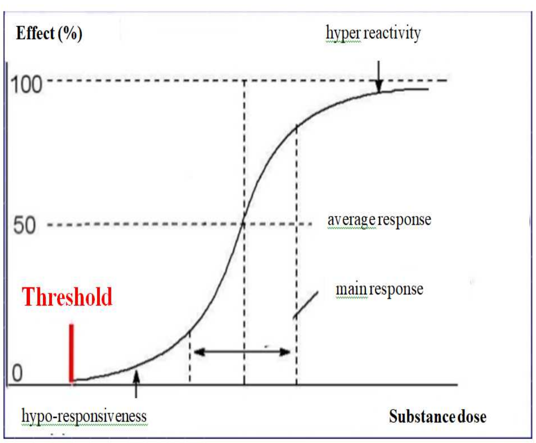


Figure 1: Dose-effect Graphic Chart. Source: Student's Course Book Introduction to Toxicological Chemistry

Hazardous effects always with liquid or food depend somewhat on the amount of contaminant or its dose in a body. The amount of a dose depends on the ways of entering tracts into the body. Contaminants can have various effects depending on their inflow type: with breathing (inhalation), water or food (orally), or absorbing through skin, or through external irradiation.

A considerable amount of research has been done in the sphere of ecology and health but little research quantitatively and qualitatively evaluates the impact of ecology on health. And yet less attention has been paid to IT-modeling the situation. As stated above, many "ecology – health" systems include the processes that are not sufficiently investigated, and which have still no formal models. Since the consequences of the changes in the ecological system or unfavorable work conditions can cause serious people's health problems, there is a great need for new knowledge. Large amounts of information are available, but analysis of large amounts of data generated by ecological systems is required, and the information remains underused. The peculiarities of environmental processes affecting public health require a new paradigm to improve analysis and, therefore, management. In order to solve environmental health security problems, people need approaches beyond the ordinary application of conventional technique. Data mining techniques provide powerful tools for extracting valuable information from large databases and are capable of identifying and capturing key parameters driving these complex systems. Currently, methods of intelligent data processing, called Data Mining, are widely used [25-29]. The term "intelligent data processing" is often translated as data mining, information extraction, mental processing, means of pattern search, pattern analysis, and knowledge search in databases.

It is necessary to conduct a survey of the possibility of predicting the number of cases when the ecological situation changes. To do this, attract modern technologies and new research in this area.

The purpose of this study is new opportunities for data mining and neural network technologies use for formalization, modeling and practical implementation of algorithms for predicting the possible number of sick people, depending on the environmental situation, using the example of Ust-Kamenogorsk city, Kazakhstan.

The objects of research are industrial enterprises and transport, the state of the air environment of the city and the health of the population of Ust-Kamenogorsk. The subject of the study is assessment of the impact of air pollution on public health.

The work used complex studies, including system analysis and generalization of scientific research results, environmental monitoring, and methods of mathematical statistics, neural nets and fuzzy algorithms and principles of consistency, complexity and mutual linking of particular approaches with general natural laws.

Sufficient data have been accumulated on the daily emissions of industrial enterprises and transport in the city, as well as the number of patients treated for the same period. The data is diverse and requires processing for analysis. The use of modern technologies will make it possible to quickly process such information and predict the possible result. The study will be useful for ecologists, doctors, people who are engaged in protecting the population from the influence of unfavorable environmental factors, scientists who deal with environmental problems and their impact on public health.

A data processing model was developed in the course of analysis of the tasks of acquisition, analyzing and controlling primary and summary data. [26]. It can be described with the graph (Figure 2).

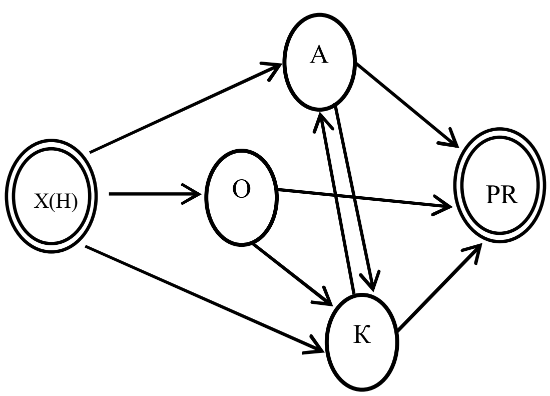


Figure 2: Data processing model

The process of accumulation (X (H)) of the control object data is a source and the decision-making process (DM) is a run-off. The arc (X (H), O) shows many instances of accounting objects and their indicators that the reporting process (O) receives from the accumulation process.

The data analysis process (A), depending on the type of data analysis task, receives a set of instances of accounting objects of the same type with selected indicators belonging to this type, or an instance of an accounting object with values a periodic indicator belonging to this instance of the accounting object, from the accumulation process (arc (X (H), A)). The data control process (K) receives many instances of accounting objects of the same type with an indicator belonging to this type from the accumulation process (arc (X (H), K)). From the progress report process, the decision-making (DM), data analysis and control processes (arcs (O, PR), (O, A) and (O, K)) receive a set of relationships between sets of control object with final indicators (obtained by aggregating), characterizing the state of the control object. From the data analysis process, the data control and decision-making processes receive (arcs (A, K) and (A, PR)), depending on the task of data analysis, many instances of accounting objects, divided into different groups, or an instance of an accounting object with of periodic indicators for future accounting periods. From the data control process, the data analysis and decision-making processes receive (arcs (K, A) and (K, PR)) for many instances of accounting objects divided into different groups. Based on the formalism of monitoring object representation the in terms of the data storage model given above, the following formalized representation of information flows can be given:

Arc (X(H),O): Vo, P, where Vo – multiple instances of accounting objects, P – set of rates.

Arc (X(H),A): Vot, Xpt1, o1, p1, Sn, Vot – multiple accounting objects of t, Xpt1 type – a subset of a set of rates associated with the type t, o1 $\in$ Vo– an instance of accounting objects with a periodic rate p with a term set Sn for n calculation periods.

Arc (X(H),K: Vot, p1, where Vot – multiple instances of accounting objects of t type, with p1.

Arc (O,PR), (O,A), (O,K): Vo, P'=f(P), where P' – the set of total indicators obtained by applying the function f to the set of indicators P. The function f can be a function of aggregation, arithmetic average, congruence addition, minimum selection, maximum selection, depending on the monitoring task.

Arcs (A,K), (A,PR) – Voz, o1, p1, Sn+k, where Voz – a set of z groups (classes, clusters, subsets) of instances of accounting objects, where o1 $\in$ Vo is an instance of an accounting object with a periodic indicator p with a set of Sn + k values for n + k accounting periods.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Arcs (K,A), (K,PR) – Vox, where Vox – group x set (subsets) of instance of a control object.

We consider the formalization of the classification problem. The classification process consists in dividing a set of accounting objects into classes according to a specific criterion. In terms of the representation model of the monitoring object, we can say that the task of classification is to determine the value of the dependent indicator of the accounting object upon the values of other indicators characterizing this object. The analyzed indicators can be any observable primary indicators in the database of the monitoring system. The object property to be assigned to one of the classes defined by the analyst can be a dependent indicator. Formally, the problem statement is as follows [27].

There are a lot of Vo accounting objects. Each oj object from a Vo set is characterized by a set of indicators of Xpj.

$$Xpj = \{p1, p2, \ldots, ph, \ldots , pm, y\}, \quad (1)$$

where ph –independent indicators which values are known and which determine a dependent indicator y. Independent indicators, the values of which are known and on the basis of which the dependent indicator is determined. The set of S values that the exponent y can take is defined and final.

$$S = \{s1, s2, \ldots , sr, \ldots, sk\}, \quad (2)$$

where k – the number of classes into which the accounting objects under study are divided.

In the classification problem, the detected functional dependence can be represented with a mathematical function. In this case, the accounting objects under study will be considered as points in (m+1) - dimensional space. Then the set of Xpj indicators of the oj accounting object is represented as coordinate, and the function is as follows:

$$yj = \omega0 + \omega1p1 + \omega1p2 + \ldots + \omega mpm, \quad (3)$$

where ω0, ω1, … , ωm – the weight of independent indicators, the classification function is to find it. In this case, all categorical and logical indicators should be converted to numeric ones.

Clustering aims at dividing the studied accounting objects into groups of “similar” objects, called clusters. In contrast to the classification problem, in the clustering problem, each of the accounting objects of the administrative monitoring system is assigned to one (or several) of the previously undefined classes. The accounting objects are divided into clusters when they are formed simultaneously. Defining clusters and dividing accounting objects into them is expressed in the final data model, which is the solution to the clustering problem.

Formally, the clustering problem is described as follows [27]. There are many Vo accounting objects under study. You need to build a set of clusters C and map F of the set Vo to the set C.

Mapping F specifies a data model which is a solution to the problem:

$$F: Vo \rightarrow C,$$
$$C = \{c1, c2, \ldots, ck, \ldots , cg\}, \quad (4)$$

where ck – a cluster containing “similar” objects from a Vos set:

$$ck = \{oj, oi \mid oj\ Vo\ \&\ oi\ Vo\ \&\ d(oj, oi) < \sigma\}, \quad (5)$$

where σ – a value determining the proximity measure for including objects in a single cluster; d(oj, oi) is a measure of proximity between objects, called distance; d(oj, oi) – a proximity measure of p between objects, called distance.

The value d(oj, oi) is called the distance between the accounting objects oj and oi, if the following conditions are fulfilled:

1 d(oj, oi) > 0, for all oj and oi.
2 d(oj, oi) = 0, if and only if oj =oi.
3 d(oj, oi) = d(oi, oj).
4 d(oj, oi) ≤ d(oj, or) + d(or, oi).

If the distance d (oj, oi) is less than some value σ, then the elements are said to be close and fit in the same cluster. Otherwise, they say that the elements are different from each other and they are referred to different clusters.

Environmental pollution can affect health in any of the following ways [30]:

- ✓ intense exposure to toxic materials or agents causing an immediate acute reaction and health consequences;
- ✓ low exposure to toxic materials or agents causing acute or chronic illness long after exposure;
- ✓ genetic changing effects;

- ✓ diminished resistance to infections;
- ✓ cause subclinical irritation, dysesthesia or discomfort;
- ✓ contribute to the exacerbation of the existing disease;
- ✓ create conditions that are incompatible or humiliating to physical, mental health and social well-being.

The air pollution exposure from different sources is difficult to clearly distinguish, and it is important to consider the exposure of mixtures of substances. These mixtures can give simply additive effect, in most cases they are often found to be synergistic, although sometimes they can be antagonistic.

The concepts of "quality of life" and "population's well-fair" are used to characterize public health: these concepts mean a certain combination of economic, social, environmental and other conditions of human life having a significant impact [31, 32]. J. Forrester proposed to measure the quality of human life by five global factors: population size, amount of accumulated capital, share of capital in agriculture, amount of available natural resources and environment contamination.

The past twenty years, according to experts' data from the city SES and the committee of public expertise, the public health in the East Kazakhstan region has deteriorated noticeably [33-35]. According to the data, the average content of lead in the human blood, for example, in Ust-Kamenogorsk, Ridder and Altai is more than 0.6 mg / l, which exceeds the threshold level, i.e. in other words, there are practically no people in Ust-Kamenogorsk with a blood lead level of less than 0.35 mg / l [36, 37]. In most areas of observation, independent public expertise found such harmful substances as vanadium, zinc, arsenic, thallium, boron, lead, chromium, and beryllium in child biological media, which exceeds content standards for these elements ranges from 1.5 to several tens of times ... For example, the content of chromium in the blood of children is 50 times higher than the natural norm in the area of Ust-Kamenogorsk, Zashchita region, and in KShT region, Ust-Kamenogorsk, the boron content for children exceeds the norm approximately 40 times [36, 37]

According to the Committee for Public Expertise's assessment of the influence of atmospheric air pollution on the level of morbidity of the population, showed that air pollution influenced the total morbidity in general, and its contribution was approximately 37% in children and about 10% of diseases in adults. Diseases of the respiratory and metabolic organs were taken into account, which amounted to 16% and 7%, respectively.

But individual constituent atmospheric pollution contributes to the development of diseases in a different way [38-46]. For example, the contribution of widespread atmospheric air pollutants (dust, carbon monoxide, sulfur dioxide, nitrogen oxides) to the development of chronic bronchitis exceeds the contribution of specific pollutants (organic compounds, metals) by about 1.6 times, while the contribution of specific pollutants is 1.5 times higher with the development of endocrine diseases. The list of substances causing immediate effects and their occurrence [41] is given in Tables 1 – 5.

The ranking of the identified emissions by the value of the total annual emissions and risk assessment for public health enabled identifying priority chemicals with carcinogenic and non-carcinogenic toxic effects, as well as the main stationary sources of atmosphere pollution. It has been established that emissions of nitrogen dioxide, sulfur dioxide, hydrogen fluoride, formaldehyde, associated with metallurgical and energy production are the greatest potential hazard.

Low indicators of the public health were detected in Ust-Kamenogorsk. The reason for such health indicators lies in the combination of the effects of the current environment impact, lifestyle indicators, working conditions, social and economic factors with the changed background caused by the intense impact of past environmental factors.

Comparative analysis of the incidence rate in Ust-Kamenogorsk, as well as the assessment of the morbidity structure and its dynamics indicates the presence of regional specific background, characterized by high levels of diseases of the blood and hematopoietic organs, genitourinary disorders, and respiratory diseases [40, 41].

Respiratory diseases are stable at a high level. The following classes of diseases have been identified with a reliable upward trend: diseases of the blood and hematopoietic organs (general population); endocrine system diseases (general population); eye disorders (children); respiratory diseases (adults, adolescents); diseases of the digestive system (general population); diseases of the skin and subcutaneous tissue (adolescents, children); genitourinary disorders(general population). The prognosis for these types of diseases is poor.

Diseases of the upper respiratory tract and respiratory organs comprise the largest share in the overall structure of morbidity - 29.5% per 10 thousand of the population. By age groups, people aged 30-40 and over 60 are more often affected by acute upper respiratory diseases. The most common are the upper respiratory tract diseases such as vasomotor and allergic rhinitis, chronic pharyngitis, chronic bronchitis, bronchial asthma e.g. diseases, caused by environmental problems of the city, main etiological factors [40, 41].

The second place takes diseases of the cardiovascular system, with 10.8% per 10 thousand of the population. Third place take diseases of the genitourinary system, with 9.4% per 10 thousand of the population.

The rest of the nosological forms have the incidence rates characterized by both periods of growth and periods of decline.

Along with this, it should be noted a dramatic fertility decline and a steady trend towards an increase in the population mortality rate, which increased by more than 10% as compared to 2000, with more than 50% of all cases from the cardiovascular system diseases and malignant neoplasms. For men, mortality is especially high from lung cancer, prostate cancer; for women from cancer of the colon, and mammary glands. Infant mortality is also very high. Congenital anomalies occupy a significant place in the structure of infant mortality [43] .

Currently, it is established that the immune status can be changed in 18-27% of patients with infectious chronic diseases. Whereas it is observed that children's immune status in Ust-Kamenogorsk has changed in 39% of conditionally healthy children. Thus, the continuing degradation of the environment in East Kazakhstan and the associated clear tendency of decreasing health indicators of the population strongly dictate the advisability of conducting comprehensive hygienic research in this region.

Thus, further degradation of the environment in East Kazakhstan and the associated steady downward trend in population health indicators strongly suggest comprehensive health studies in this region. The above study makes us to conclude that the decline in public health is influenced by the deteriorated environmental conditions. In recent years, there has been a tendency for the deterioration of the public health status in all groups of morbidity.

The growth of oncological and other diseases and the inevitable payback for the mismanagement of nature, are due to environmental laws, according to which, any change in the environment will inevitably cause progressing natural chain reactions resulting in neutralization of the changes made. This interaction is not linear, that is, even a weak impact on one of the indicators can cause reciprocal strong deviations in others, as well as in the entire human-nature system as a whole, which is observed in the East Kazakhstan region.

A correlation analysis was carried out to solve the problems of assessing and predicting the dose-effect of the impact of complex environmental pollution on the public health indicators, which enables a quantitative relationship ratio between the concentration of pollutants and indicators of public health, assessment the strength and nature of the air pollutant influence on the morbidity and mortality of the population.

On the basis of the studied literature [36, 37, 40, 41], an analysis of the state of fertility and mortality in the region of the East Kazakhstan region was carried out.

Since 2000, the birth rate in the East Kazakhstan region has been decreasing through 2012, inclusive, from 33794 to 16306 people. Recently, the birth rate has increased by 3,332 people and amounted to 19,638 people in 2015.

Life expectancy at birth in the East Kazakhstan region since 2012 comprises 65.37 years, it decreased by 2015 by 0.59 years, which amounted to 64.78 years. Moreover, the urban population has a significantly lower life expectancy at birth than the rural population.

The current demographic situation in Kazakhstan is characterized by a steady dynamic increase of older people, which corresponds to the global aging process of the population. The aging process of the population, expressed in the increase in the older people proportion, has been caused by a decrease in the birth rate and an increase in the total life expectancy at birth; nowadays, it is observed almost globally. According to UN forecasts, by 2050 the number of elderly people will be about 1.5 billion, which corresponds to 14.7% of the total population which will result in a decline in the working-age population. Population aging is the result of two factors: increased life expectancy and high mortality rates, especially among men of working age.

Mortality rate is one of the main criteria for assessing and characterizing public health status and its changes. Although mortality is a phenomenon of the natural development of life, it is socially conditioned and depends on many factors, which are divided into two large groups: endogenous (resulted from the internal development of the human body) and exogenous (related to the external environment). An unfavorable environmental situation significantly affects the health of the entire population. These factors are air and water pollution. There is also Semipalatinsk nuclear test site in East Kazakhstan region, due to which the soil is contaminated. Large enterprises in Ust-Kamenogorsk such as Ust-Kamenogorsk Titanium-Magnesium Plant, Kazzinc JSC, Vostokmashzavod JSC (AsiaAvto), Ulba Metallurgical Plant JSC, Ust-Kamenogorsk TPP, Sogrinskaya TPP, pollute the environment of our area. Ulba Metallurgical Plant has started a large cycle of investment "innovation" program, dealing with three main industries of this enterprise, namely: uranium, tantalum and beryllium, which will negatively affect the health of the population, which is 20.9% of the region.

3. MODELS AND ALGORITHMS

The task of classification or pattern recognition can be solved using various neural network topologies. Multilayer neural perceptions, classical networks for classification, due to their approximation qualities, remain the leading tool among neural networks for supervised classification problems in medium-sized spaces. Convolutional neural networks have a powerful potential for pattern recognition under conditions of information distortion and high dimensionality of input and output sets (for example, recognition of cartographic images and handwritten texts), but to use them for solving simpler problems with a lower dimension is not entirely rational, since their implementation is more time consuming, training is more protracted, and visual explanation of the learning process is lost. Thus, a multilayer perceptron is the most suitable neural network topology for the problem of supervised classification of situations of the influence of environmental factors on the public health.

The general algorithm for training a network can be represented as follows:

1) Initialization of network weight with randomized small values.

2) Entering the training sample vector X to the network input and calculating the NET signal from each neuron.

$$NET_j = \sum_{i=1}^{n} w_{ij} x_i \quad (6).$$

3) Calculation of the value of the threshold functions of the neuron for the NET signal for each neuron:

$$OUT_j = \begin{cases} 1, NET_j > \theta_j \\ 0, NET_j \le \theta_j \end{cases} \quad (7),$$

where θ j is the threshold corresponding to the j-th neuron (an example for the simplest version of the threshold function).

4) The error is calculated for each neuron:

$$e_j = Tr_j - OUT_j \quad (8),$$

where Tr_j –a required neuron output

5) Neuron weight modification:

$$w_{ij}(t+1) = w_{ij}(t) + \alpha x_i e_j \quad (9),$$

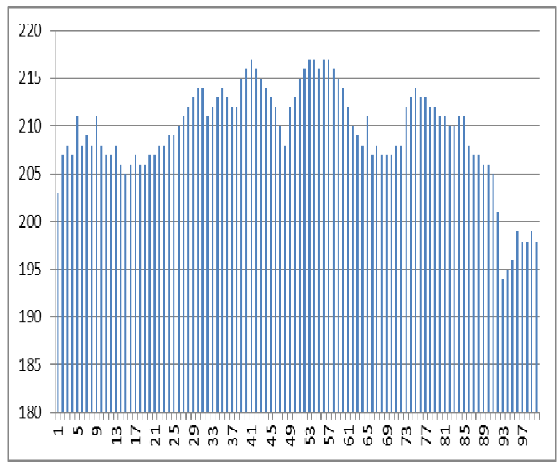
where α - a coefficient that provides adjusting the network learning rate. The problem of training multilayer network is essentially reduced to the problem of multi-criteria unconstrained optimization. To solve it such methods as Newton's optimization method, steepest descent methods, including gradient methods, and their specialized neural network modifications, such as error backpropagation algorithm, are used.

The successful training of a neural network can decisively depend on the form of information presented for training. Data normalization is the first stage of preprocessing, therefore it is important for data preprocessing for a neural network. The subsequent process of training and analysis of the neural network functioning will depend on the results of correct data normalization. Normalization is considered to be the adjustment of a series (vector) of values in accordance with some transformation functions in order to make them more suitable for analysis and (or) comparison. Data normalization is required when the incompatibility of units of measurement of variables can affect the results (as in the case of training a neural network), and is recommended in cases where the final reports can be improved by expressing the results in certain understandable / compatible units.

So, heterogeneous data will be fed to the inputs of the designed neural network in order for them to have equal weight during training of the neural network and make an equal contribution to the error function. It is necessary to normalize these data. It is important that if, from the point of view of the multiplication operation, the values of ±1 are equal, there is a significant asymmetry between 0 and 1: zero values do not make any contribution to the error gradient. Thus, the choice of the input coding scheme influences the learning process. Due to the logical equality of both values of the inputs, the symmetric encoding is more preferable: $\{-1;1\}$, which preserves this equality in the learning process. To normalize the bulk of the data and simultaneously limit the range of possible values, sigmoidal neuron activation functions can be used for data preprocessing. For example, nonlinear transformations using the hyperbolic tangent function:

$$\tilde{x}_i = th\left(\frac{y}{\alpha}\right) \qquad (10),$$

where $y = \dfrac{x_i - \bar{x}_i}{\sigma_i}$, and α - the slope coefficient of the sigmoidal function, which normalizes the bulk of the data and enables placing the data in the interval $\tilde{x}_i \in [-1;1]$. This prevents the danger of asymmetry and uneven contribution of inputs when calculating the gradient during training, which can happen when normalizing using the logistic function, since applying it all data is placed in the unit interval $\tilde{x}_i \in [0;1]$. To select the most optimal normalization method, we carry out an experimental normalization of the weighted average harmful emissions into the atmosphere and the corresponding increase in diseases associated with neoplasms (Figure 3), which we prepare to feed to the neural network inputs. These data are heterogeneous, since the average emissions vary in the range from 194 to 217, and the indicators from 0 to 100 points.



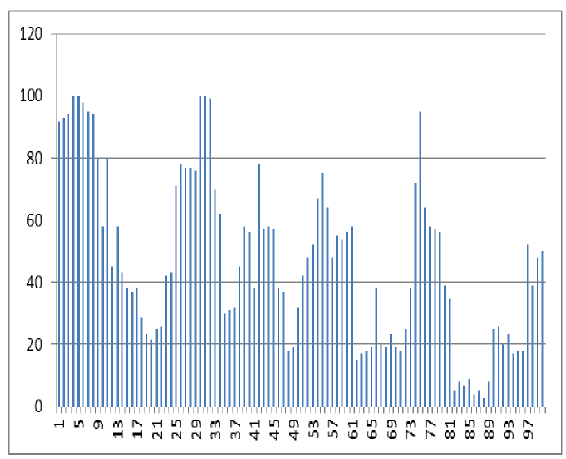


Figure 3: Daily averaged emission value (top) and the indicator of the diseases growth (bottom) from 02/01/2015 to 02/16/2015

Normalization using the activation function of neurons of the hyperbolic tangent provides for the slope coefficient selection. With a slope factor α = 1, spikes are found in the data (Figure 4).

Journal of Theoretical and Applied Information Technology
15th August 2021. Vol.99. No 15
© 2021 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

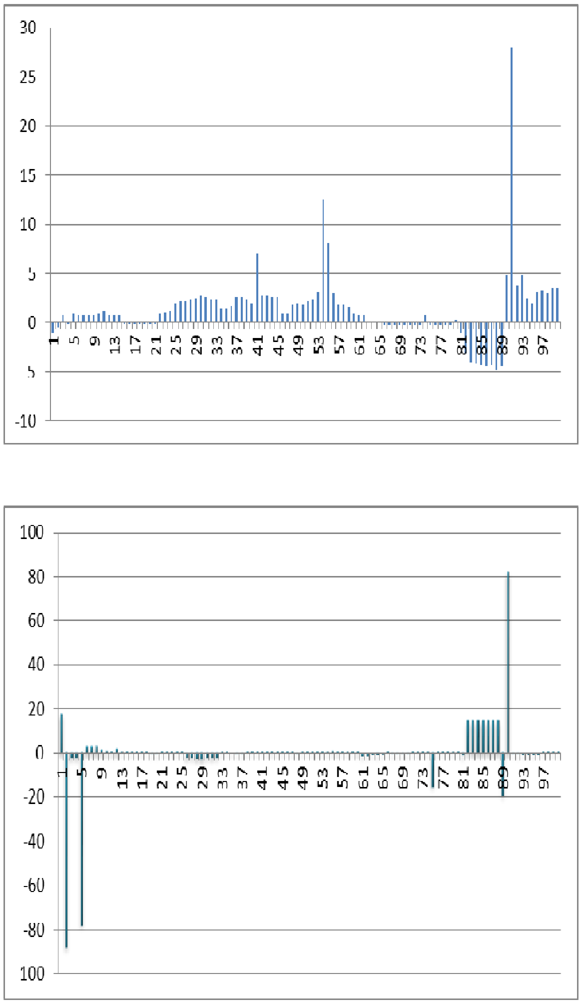


Figure 4: Nonlinear normalization of the mean value of emissions (top) and the indicator of the increase in diseases (bottom) using the hyperbolic tangent (slope factor = 1)

To eliminate surges and bring all data into the interval $\tilde{x}_i \in [-1;1]$, the following algorithm must be applied:

- ✓ find out the maximum modulus value from the data with normalized by a hyperbolic tangent with a slope coefficient of 1;
- ✓ find out for this value the value of y from the formula y_{max};
- ✓ -calculate the value of the slope factor by the formula:

$$\alpha = \frac{y_{\max} \cdot 4}{\pi} \qquad (11),$$

based on $tg(\frac{\pi}{4}) = 1$.

Thus, we obtain samples normalized relative to the mean values from heterogeneous data and both groups of data are located in the interval [−1;1] (Figure 5), which increases the informative data component, that is the information entropy.

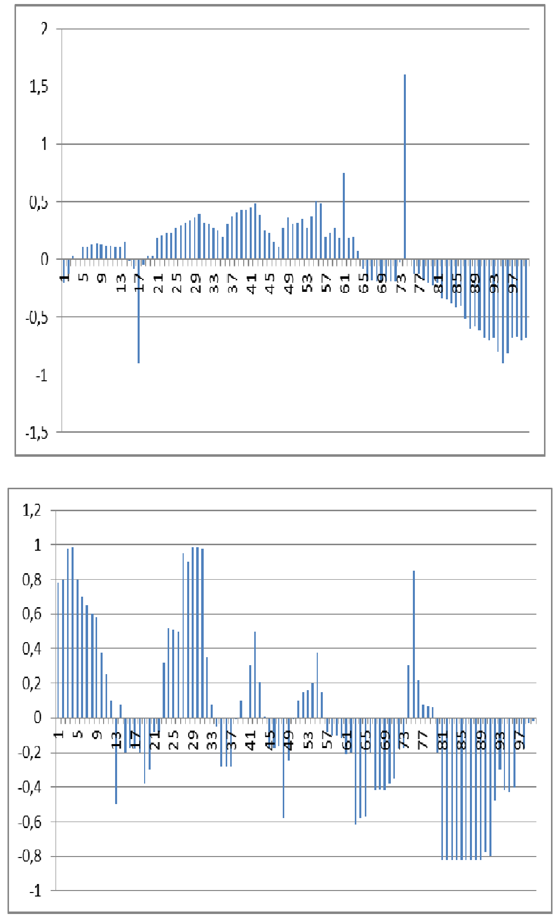


Figure 5: Nonlinear rationing of the mean emission value (top) and disease increase indicator (bottom) with a hyperbolic tangent with a well-defined slope values

The averaged values of emissions and the rate of the disease increase, thus normalized, become valid for neural network analysis. They will effectively train a neural network to classify environmental situations. Each new datum of both the averaged emission value and the rate will equally contribute to the learning process [25].

To train a neural network means to supply it with initial data and the results of the analysis, according to which it can build the necessary relationships for a similar analysis based on similar data (Figure 6).

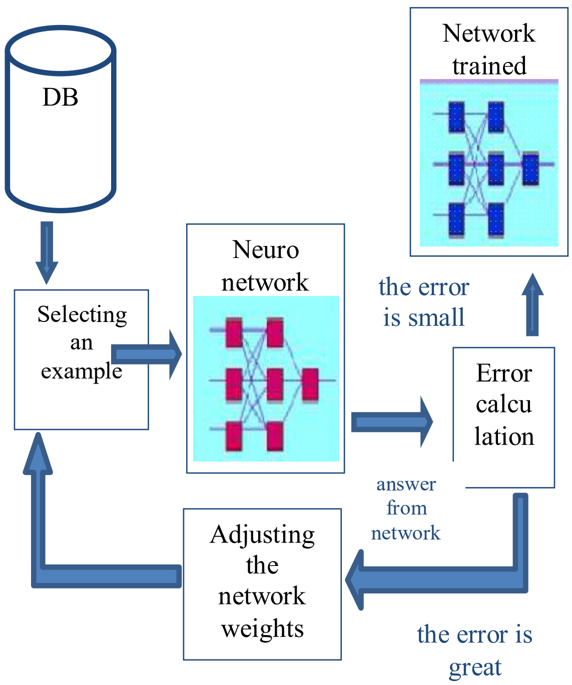


Figure 6: Conceptual model of neural network learning process

A model of the impact of harmful emissions on the public health in Ust-Kamenogorsk was built to test the acceptable neural network technologies for assessing environmental risk. As an environment state indicator, the air pollutant concentration was used according to the data of the Kazhydromet Center for Environmental Pollution Monitoring, which monitors the quality of atmospheric air at permanent sampling stations in Ust-Kamenogorsk (Tables 1 – 5). The yearbooks "The state of air pollution in cities on the territory of East Kazakhstan region" from 2005 to 2015 were processed [41].

The data of sanitary and demographic statistics were used as an indicator of population health. To build a neural network model, the Matlab software system (version 7.12.0.635) was used. The training of the neural network was carried out with the data for 2005-2015, while testing with the data for 2014-2015, and forecasting according to data for 2011.

The artificial neural network training process corresponds to the conceptual model of the neural network training process presented above in Figure 6.

Table 1: Pollutant concentration 1-5

Year	Substance, mg/m^3				
	benzol	Nitrogen oxid	marganese	opper	Hydrofluoride
	1	2	3	4	5
2005	0.081	0.020	0.041	0.03	0.0031
2006	0.029	0.019	0.035	0.02	0.0030
2007	0.023	0.027	0.029	0.03	0.031
2008	0.015	0.028	0.040	0.04	0.030
2009	0026	0.018	0.041	0.03	0.029
2010	0.014	0.039	0.080	0.04	0.0010
2011	0.023	0.034	0.070	0.05	0.0021
2012	0.018	0.041	0.100	0.04	0.0030
2013	0.024	0.048	0.090	0.06	0.0280
2014	0.020	0.062	0.091	0.05	0.025
2015	0.025	0.058	0.081	0.09	0.0018

Table 2: Pollutant concentration 6-10

Year	Substance, mg/m^3				
	Xylene	Hydrogen sulfide	Carbon disulfide	Toluene	carbon monoxide
	6	7	8	9	10
2005	0.051	1.0	5.000	0.045	1.500
2006	0.070	0.9	1.000	0.031	1.400
2007	0.0014	0.8	0.002	0.014	1.800
2008	0.013	0.01	0.002	0.013	1.600
2009	0.018	0.02	0.001	0.029	1.400
2010	0.028	0.01	0.300	0.025	1.450
2011	0.042	0.3	0.100	0.031	1.075
2012	0.036	0.05	0.100	0.027	1.800
2013	0.053	0.2	0.300	0.038	0.890
2014	0.045	0.3	0.250	0.040	1.190
2015	0.029	0.2	0.225	0.025	1.200

Several network models were designed by varying the structure and parameters of the neural network and the learning algorithm. The best one was selected by the generalization property (the smallest error on the test sample). The selected network has a layer-by-layer structure and direct signal propagation (multilayer perceptron) with three layers: layer 1 - 3 neurons, layers 2 and 3 - 2 neurons each with a sigmoidal activation neuron function. The average relative error for all results was 0.40%, the average absolute error was 0.93%, and i.e. the network provides good convergence of calculated and actual values.

Table 3: Pollutant concentration 11-15

Year	Substance ,mg/m^3				
	Phenol	Nickel	Lead	Benz pyrene	Chromium
	11	12	13	14	15
2005	1.0	0.0125	0.02	2.10	0.0225
2006	1.1	0.0113	0.01	2.11	0.0212
2007	1.1	0.0101	0.01	3.90	0.0201
2008	0.9	0.0100	0.02	2.78	0.0125
2009	0.8	0.0200	0.02	3.65	0.0020
2010	1	0.0200	0.02	4.08	0.0221
2011	1.1	0.0200	0.04	3.35	0.0300
2012	0.8	0.0211	0.04	3.41	0.0222
2013	1.2	0.0210	0.07	4.41	0.0210
2014	1.1	0.0100	0.12	4.80	0.0210
2015	1.5	0.0100	0.17	4.98	0.0211

Table 4: Pollutant concentration 16-20

Year	Substance, mg/m^3				
	Formalde-hyde	Nitrogen dioxid	Zinc	Dust	Ammonia
	16	17	18	19	20
2005	4.8	0.026	0.078	0.32	0.020
2006	4.9	0.025	0.073	0.25	0.018
2007	7.8	0.024	0.070	0.27	0.017
2008	7.9	0.035	0.073	0.26	0.020
2009	8.0	0.040	0.070	0.27	0.018
2010	5.8	0.038	0.089	0.20	0.022
2011	11	0.045	0.051	0.23	0.024
2012	5.5	0.051	0.080	0.23	0.018
2013	6.5	0.059	0.071	0.22	0.019
2014	8.9	0.062	0.071	0.19	0.030
2015	13	0.070	0.078	0.18	0.031

Table 5: Pollutant concentration 21-25

Year	Substance, mg/m^3				
	Sulfur dioxide	Iron	Chlorine	Hydrochloride	Ethylbenzene
	21	22	23	24	25
2005	0.0019	0.42	0.030	0.043	5.78
2006	0.0020	0.78	0.030	0.041	9.78
2007	0.0040	1.15	0.014	0.040	10.1
2008	0.0039	1.50	0.008	0.045	12.1
2009	0.0050	1.51	0.007	0.048	12.7
2010	0.0048	2.01	0.006	0.050	13.0
2011	0.0045	3.01	0.001	0.0057	12.8
2012	0.0037	3.51	0.001	0.051	14.2
2013	0.0024	4.00	0.002	0.068	15.2
2014	0.0025	3.55	0.001	0.070	15.0
2015	0.0023	7.38	0.001	0.050	15.1

The forecast obtained for the number of deaths from neoplasms in 2011 - 199.06 per 1000 people, while the real number is 199.08.

Calculations based on the proposed approach, apparently, have a greater degree of uncertainty compared to the results obtained with exact requirements of scientific biomedical expertise. However, its clear advantage is that now there are all the necessary legal, methodological and informational opportunities for its implementation in the practice of ecological expertise and control over natural environment condition. In addition, the proposed method of risk analysis an enable, if necessary, to obtain objective quantitative information about the degree of danger of a specific operating or projected production facility, to identify zones and territories with the level of risk exceeding the permissible values, at which it is necessary to tighten control or take certain measures to reduce it and ensure the regulatory safety of production staff and population [46-54].

An artificial neural network (ANN) is a mathematical apparatus that allows to build processing algorithms, which has a unique ability to learn by examples and "recognize" in the stream of "noisy" and contradictory information, signs of previously encountered images and situations. ANN allows finding hidden dependencies between input and output data, which are beyond the attention of classical methods [25-27].

The above results indicate that the use of neural network technologies for solving applied environmental problems dealing with information processing and building models is a promising area of research.

Based on the foregoing, the developed intelligent information system for assessing the negative impact of industrial emissions on the public health (using the data from Ust-Kamenogorsk as an example) allows to predict unfavorable environmental situations and to consider the possible scenarios.

4. PROGRAM IMPLEMENTATION

Main specifications:

- ✓ Microsoft Windows 10 any version,
- ✓ processor x64, x86.
- ✓ screen resolution not less than 1536 by 864 pixels at 100% scale.

The software product is started by clicking on the EnvironmentalDataAnalysis icon.

There are five main tabs of the software available:
1) Environmental posts;
2) Environmental monitoring;
3) Weather monitoring;
4) Morbidity data;
5) Neural network and forecasting.

The Environmental posts tab allows to download actual data from the posts of environmental monitoring of the city of Ust-Kamenogorsk. The posts monitor the following harmful components:

- ✓ carbon dioxide CO_2;
- ✓ carbon monoxide CO;
- ✓ hydrocarbons CxHy;
- ✓ chlorine hydride HCl;
- ✓ formaldehyde HCOH;
- ✓ hydrogen fluoride HF;
- ✓ nitrogen dioxide NO_2;
- ✓ sulfur dioxide SO_2.

Information from 9 ecological posts is avaliable: Promzona, Zashchita station; UK Station; KShT; Pontoon bridge; Mirny settlement; Obyezdnaya Road; Sogra; Abay Avenue.

The Environmental Monitoring tab allows you to load real ecological monitoring data and design air pollution graphs.

The date and time of the measurement, the post of environmental monitoring, as well as the content of harmful components are displayed in the left part of the window (Figure 7).

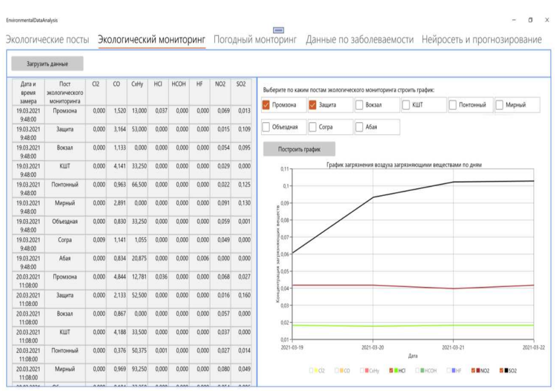


Figure 7: Environmental monitoring data window: visualization of data of two environmental monitoring posts on three hazardous components

On the right part of the window, you can choose hazardous substances and environmental monitoring posts to build a graph by setting flags. The information is visualized by clicking on the Build a graph button.

The Weather Monitoring tab displays information by the Kazhydromed meteorological service, including date, minimum and maximum air temperature, atmosphere pressure, wind direction and speed (Figure 8).On the left side of the window, there is weather monitoring data, and on the right side of the window there is a graph of temperature fluctuations by day (Figure 8).

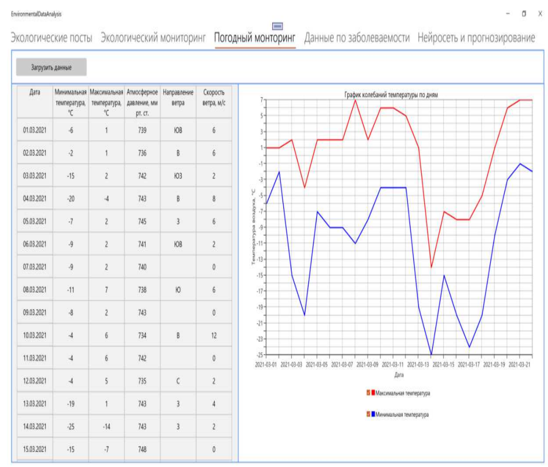


Figure 8: Window with a graph of daily temperature fluctuations

The tab Morbidity Data provides information on the date, the number of cases in the context of adults and children, and the total number of cases. The data on acute respiratory viral infections with symptoms of respiratory tract disease are taken as an example.

The methods of forecasting, building and training a neural network described in the previous chapters are represented on the “Neural Network and Forecasting” tab (Figure 9).

The process of model training is carried out in several stages. At the first stage, averaged data on pollutants are entered to the inputs of the model at the environmental monitoring posts selected for drawing the graph. Model outputs can optionally be the number of adults, children, or the total number of cases. When you click on the button Prepare dataset for training in the window Event Log, an alert appears: “Preparation of dataset for training has started”. The data processing uses the algorithms described in the previous chapters. At the end of the preparation of the dataset, the corresponding notification appears in the event log window: "The preparation of the dataset for training is finished" and the number of loaded records. To start the second stage of model training it is required to press the button "Train the model", and the model metrics can be calculated to assess the proximity and adequacy of the modeling.

For forecasting, the data on pollutant content can be entered manually (Figure 9).

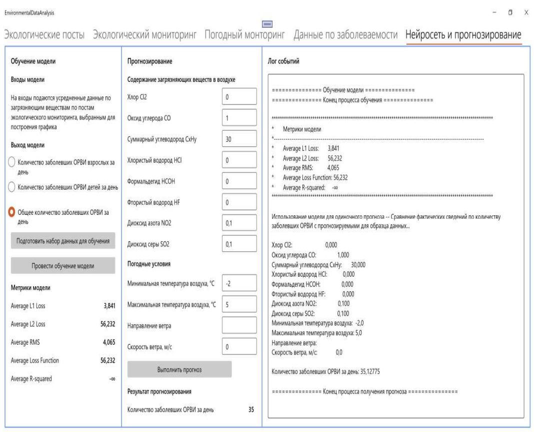


Figure 9: Forecasting window

After entering all the forecast data, you should click on the "Run forecast" button. The event log window will display data on the use of the model and the comparison of the factual data on the number of cases, for example, SARS with the predicted data for the sample.

Figure 10: A certificate on data registration in the state register of copyrighted objects of the Republic of Kazakhstan No. 16777, from April 20, 2021

The prediction result is displayed in the corresponding Forecasting window and in the event log.

All the considered methods and algorithms are implemented in the software package in C# programming language, providing visualization of forecasting results.

5. ACKNOWLEDGEMENT, CONCLUSION AND FUTURE RESEARCH DIRECTIONS

Part of the research was carried out with the support of the "Sustainable Mobility Research in Central Asia (SUMRICA) project" and Technische Universität Berlin.

The goal of the work has been achieved; the research tasks have been completely solved. The research results have been brought up to implementation, which confirms the reliability of the main provisions and conclusions.

It was conducted a survey of the possibility of predicting the number of cases when the ecological situation changes with the help of attraction modern technologies and new research in this area such as data mining and neural network technologies used for formalization, modeling and practical implementation of algorithms for predicting the possible number of sick people, depending on the environmental situation, using the example of Ust-Kamenogorsk city, Kazakhstan.

It is planned to continue further studies of the impact of harmful emissions on public health, taking into account new data and the introduction of new software for air monitoring in Ust-Kamenogorsk. Predicting a surge in morbidity depending on weather conditions and changes in the concentration of harmful substances will allow for operational regulation of the emissions of enterprises into the atmosphere. New methods will be developed to regulate air pollution.

REFRENCES:

[1] K.A. Vogt et al. "Ecosystem Concept: Historical and Present Review of Definitions and Development of Ecosystem Ecology, Ecosystem Management, and Its Legal Framework", *Ecosystems. Springer*, New York, NY (USA), 1997, https://doi.org/10.1007/978-1-4612-1908-8_2, accessed 14.05.2021.

[2] http://www.who.int/topics/air_pollution/en/ World-Health-Organization, Air Pollution, accessed 14.05.2021.

[3] http: // www.who.int/phe/ health topics/outdoorair/databases/health impacts/en/, World Health Organization, "Public health, environmental and social determinants of health", 2015, accessed 12.05.2021.

[3] T. Hirsch, S. Weiland, E. von Mutius, A. Safeca, H. Grafe, E. Csaplovics, H. Duhme, U. Keil, and W. Leupold, "Inner city air pollution and respiratory health and atopy in children", *European Respiratory Journal,* 14, 1999, pp. 669-677.

[4] J. Park, S. Mackay, "Practical Data Communications for Instrumentation and Control", *Elsevier science & technology,* Oxford (United Kingdom), 2003, 400p.

[5] V.V. Smorodin, E.V. Volkova, A.A. Aliev "Ot khraneniia dannykh k upravleniiu informatsiei", ISBN: 978-5-4237-0008-9 *SPb: Piter*, 2010, 528 p.

[6] M.E. Berliand, "Prognoz i regulirovanie zagriazneniia atmosfery", Gidrometeoizdat, 1985, 272 p.

[7] E.A. Vukolov "Osnovy statisticheskogo analiza", *FORUM 2008,* Moscow (Russia), 464 p.

[8] Y. Yang, H. Sun, J. Xue, et al., "Estimating evapotranspiration by coupling Bayesian model averaging methods with machine learning algorithms", *Environ Monit Assess* 193, 156 p., 2021, https://doi.org/10.1007/s10661-021-08934-1

[9] M. Power, S. Adams, "Perspectives of the Scientific Community on the Status of Ecological Risk Assessment", *Environmental Management* 21, 1997, pp. 803–830, https://doi.org/10.1007/ s002679900069, accessed 14.05.2021.

[10] A. Davis, "Environment Management", *Mastering Salesforce DevOps*. Apress, Berkeley, CA. (USA), 2019, https://doi.org/10.1007/978-1-4842-5473-8_6, accessed 21.05.2021.

[11] G.M. Mutanov, D.N. Cherkasov, "Priniatie reshenii po likvidatsii chrezvychainykh situatsii s ispolzovaniem neironnykh setei", GIAB, №2, 2001, https://cyberleninka.ru/article/n/prinyatie-resheniy-po-likvidatsii-chrezvychaynyh-situatsiy-s-ispolzovaniem-neyronnyh-setey (accessed 13.05.2021).

[12] N.M. Temirbekov, F.N. Abdoldina, M.N. Madiiarov "Ekologicheskii monitoring vozdeistviia avtotransporta na atmosferu goroda", *Vychislitelnye i informatsionnye tekhnologii v nauke, tekhnike i obrazovanii*, 7-9 Oct. 2004, Almaty – Novosibirsk, 2004, pp. 122-124.

[13] T.G. Balova, M.T. Ipalakova, "Statisticheskie modeli i metody prognozirovaniia urovnia zagriazneniia vozdukha", *Vestnik D. Serikbayev East-Kazakhstan state technical University (Kazakhstan)*, №4, 2003, pp. 20-26.

[14] O.Ya. Shvets, "Razrabotka informatsionnoi sistemy vyiavleniia chrezvychainykh situatsii. Monitoring atmosfery gorodov", Saarbrücken, *LAP LAMBERT Academic Publishing* (Germany), ISBN: 978-3-569-24063-8, 2013, 159p.

[15] U.M. Sultangazin, "Mathematical problems connected with construction of algorithms for atmosphere correction in remote sensing", *Journal of Inverse and Ill-posed Problems*, vol. 9, no. 6, 2001, pp. 655-668, https://doi.org/10.1515/jiip.2001.9.6.655, accessed 14.05.2021.

[16] B.T. Zhumagulov, U.S. Abdibekov, O.A. Akhmetov, "Turbulentnyi perenos primesi v iskrivlennykh kanalakh" *Vestn. Nats. inzh. akad. RK*, №1, 2006, pp. 24 - 27.

[17] T.Z. Muldashev, A.I. Lyapustin, U.M. Sultangazin, Spherical harmonics method in the problem of radiative transfer in the atmosphere-surface system, *Journal of Quantitative Spectroscopy and Radiative Transfer*, Volume 61, Issue 3, 1999, pp. 393-404, ISSN 0022-4073, https://doi.org/10.1016/S0022-4073(98)00025-9, accessed 14.05.2021.

[18] O.L. Zhizhimov, A.M. Fedotov, Iu.I. Shokin, "Tekhnologicheskaia platforma massovoi integratsii geterogennykh dannykh", *Vestnik Novosibirskogo gosudarstvennogo universiteta, Seriia: Informatsionnye tekhnologii*, T. 11, №1, 2013, pp. 24-41.

[19] K.T. Protasov, B.D. Belan, Z.F. Idrisov "Adaptivnaia identifikatsiia optiko-meteorologicheskikh kharakteristik atmosfery", *Optika atmosfery*, T. 1, № 10,. 1988, pp. 95-98.

[20] L.F. Nozhenkova, V.V. Nicheporchuk, "A system for distributed collecting and analysing the Emergency monitoring data", *Informatizatsiia i sviaz*, № 5, ISSN 20788320, Moscow (Russia), pp. 45-50.

[21] K.S. Baker, G.C. Bowker "Information ecology: open system environment for data, memories, and knowing". *J Intell Inf Syst* 29(1), 2007, pp. 127–144.

[22] D.G. Green, N.I. Klomp, G. Rimmington, S. Sadedin "Digital Ecology: New Technologies Are Revolutionizing Ecology". *Complexity in Landscape Ecology*. Landscape Series, vol 22. Springer, Cham, 2020, https://doi.org/10.1007/978-3-030-46773-9_10, accessed 21.05.2021.

[23] A.A. Bogdanov, V.D. Shlendov "A method of analyzing of the "Dose-effect" dependence in multifactor medico-biological experiments", *Bull Exp Biol Med* 117, 1994, pp. 334–336, https://doi.org/10.1007/BF02444181, accessed 21.05.2021.

[24] O.Ya. Shvets, "Formirovanie chrezvychainoi situatsii v usloviiakh statsionarnoi raboty predpriiatii v ramkakh ekologicheskogo monitoringa", *Nauchno-prakticheskii zhurnal «Sovremennye materialy, tekhnika i tekhnologii»*, (Russia), №4(7), 2016, pp. 185-194.

[25] T. Hastie, R. Tibshirani, J. Friedman, "The Elements of Statistical Learning: Data Mining: Inference and Prediction", *Springer Verlag*, 2001, 745 p.

[26] M. Kantardzic, "Data Mining: Concepts, Models, Methods, and Algorithms", ISBN: 0471228524, *Wiley-IEEE Press*, 2002, 672 p.

[27] C. McCue, "Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis", Second Edition, *Elsevier*, 2015, https://doi.org/10.1016/C2013-0-00434-3, accessed 21.05.2021.

[28] J. Chen, H. He, G. Williams, and H. Jin. "Temporal sequence associations for rare events, Advances in Knowledge Discovery and Data Mining", *Proceedings of 8th Pacific-Asia Conference (PAKDD 2004),* Sydney, Australia, Springer, 2004, volume 3056 of Lecture Notes in Computer Science, May 26-28,2004, pp 239–239.

[29] C. Luan, G. Dong, "Experimental identification of hard data sets for classification and feature selection methods with insights on method selection", *arXivpreprint arXiv*:1703.08283, 2017.

[30] E. Giroult, The health component of the environmental impact assessment, *International seminar on environmental impact assessment*, University of Aberdeen (Scotland), 1984.

[31] Jay W. Forrester, "Urban Dynamics", *Pegasus Communications, Inc.*, 1969, 299p.

[32] S.E. Jorgensen, G. Bendoricchio, "Fundamentals of Ecological Modelling". 3rd ed. Amsterdam: *Elsevier*. 2001, 544 p.

[33] M. Kulzhanov, B. Rechel, "Kazakhstan: Health system review", *Health Systems in Transition*, 2007; 9(7): 1–158.

[34] http://www.stat.kz/, 2013, Agency of Statistics, Republic of Kazakhstan, accessed 15.01.2021.

[35] https://www.primeminister.kz/en/news/ reviews /review -of-kazakhstans-healthcare- system-Results-of-2020-and-plans-for-20215968, 01.03.2021, 15:14, Review of Kazakhstan's healthcare system: Results of 2020 and plans for 2021, accessed 12.04.2021.

[36] L.G. Saveleva, R.A. Garapova, "Problema sezonnoi periodichnosti v strukture zabolevaemosti naseleniia g. Ust-Kamenogorska", *Ekolog. faktory i probl. prakt. Meditsiny, Sb. nauch. tr., "Media-Alians"*, Ust-Kamenogorsk (Kazakhstan), 2003, C.3- 10.

[37] L.G. Saveleva, R.A. Garapova "Klimatopogodnye usloviia i sezonnost zabolevaemosti naseleniia g. Ust-Kamenogorska", *Meditsina*, N3, 2003, pp. 40-41.

[38] Avdeenko N.V., Efimova A.A., Balobolkin I.I. i dr. "Vliianie zagriazneniia vozdushnoi sredy na rasprostranennost i techenie allergicheskikh boleznei u detei", *Pediatriia*, №5, 1999, pp.15 – 24.

[39] E.M. Aksel, M.I. Davydov, T.I. Ushakova "Statistika raka legkogo, zheludka i pishchevoda: sostoianie onkologicheskoi pomoshchi, zabolevaemost i smertnost", *Vestnik RAMN*, №9, 2001.

[40] "Demograficheskii ezhegodnik Kazakhstana", *Agentstvo Respubliki Kazakhstan po statistike*, Astana (Kazakhstan), 2005 – 2015.

[41] "Ekologicheskii monitoring sostoianiia okruzhaiushchei sredy Vostochno-

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Kazakhstanskoi oblasti”, *Spravochno-informatsionnyi vestnik*, 2005 – 2015.

[42] S. Novikova, N. Valitova, “Environmental Emergencies Prediction Based on Time Series Measurements of Pollutants Concentration in the Atmosphere”, *Actual Problems of Computer Science*, Lublin, (Poland), ECCC Foundation, No. 1(3), 2013, pp.7-18, ISSN 2299-8667.

[43] R. McConnell, T. Islam, K. Shankardass, M. Jerrett, F. Lurmann, F. Gilliland, J. Gauderman, E. Avol, N. K. Aijnzli, L. Yao, J. Peters, and K. Berhane. Childhood incident asthma and trace-related air pollution at home and school. *Environmental Health Perspectives,* 118, 2010, pp. 1021-1026.

[44] https://www.who.int/health-topics/chemical - safety#tab=tab_1, World health organization, Chemical Safety, accessed 18.01.2021.

[45] http://www.cdc.gov/niosh/homepage.html, The National Institute for Occupational Safety and Health (NIOSH), accessed 18.01.2021.

[46] FW.J Lipfert. Long-term associations of morbidity with air pollution: A catalog and synthesis. *Air Waste Manag Assoc*. 2018 Jan; 68(1):12-28. doi: 10.1080/10962247.2017.1349010. Epub 2017 Nov 2. PMID: 28679072, accessed 03.02.2021.

[47] P.H. Bakhtiari, M.R. Nikoo, F. Golkar et al., “Design of a high-coverage ground-based CO_2 monitoring layout using a novel information theory-based optimization model”, *Environ Monit Assess* 193, 2021, p.150, https://doi.org/10.1007/s10661-021-08933-2

[48] I. Manisalidis, E. Stavropoulou, A. Stavropoulos, E. Bezirtzoglou. Environmental and Health Impacts of Air Pollution: A Review. *Front Public Health*. 2020 Feb 20; 8:14. doi: 10.3389/fpubh.2020.00014. PMID: 32154200; PMCID: PMC7044178, accessed 18.01.2021.

[49] R. Sindhwani, “Assessment of gaseous and respirable suspended Particulate matter (PM10) emission estimates over megacity Delhi: past trends and future scenario (2000-2020)”, *Center for atmospheric sciences*, November 2012, pp. 123-140.

[50] L. Grout, S. Hales, N. French, M.G. Baker. A Review of Methods for Assessing the Environmental Health Impacts of an Agricultural System. *Int J Environ Res Public Health*. 2018; 15(7):1315. Published 2018 Jun 23. doi:10.3390/ijerph15071315, accessed 15.03.2021.

[51] Hasheel Tularam, Lisa F. Ramsay, Sheena Muttoo, Bert Brunekreef, Kees Meliefste, Kees de Hoogh, Rajen N. Naidoo, A hybrid air pollution / land use regression model for predicting air pollution concentrations in Durban, South Africa, *Environmental Pollution*, Volume 274, 2021, 116513, ISSN 0269-7491, https://doi.org/10.1016/j.envpol.2021.116513, accessed 02.05.2021.

[52] F. Cifuentes, A. Gálvez, C. González, M. Orozco-Alzate, B.H. Aristizábal, (2021). Hourly Ozone and PM2.5 Prediction Using Meteorological Data – Alternatives for Cities with Limited Pollutant Information. *Aerosol Air Qual.* Res. https://doi.org/10.4209/aaqr.200471, accessed 05.06.2021.

[53] Miller, C.A. Fifty years of EPA science for air quality management and control. *Environmental Management.* 67, 1017–1028 (2021). https://doi.org/10.1007/s00267-021-01468-9, accessed 20.04.2021.

[54] J. Deane Sensenbaugh. Air & Waste Management Association (2021) *Environmental Technology Award*. Air & Waste Management Association, https://www.awma.org/content.asp?contentid=403, accessed 24 Mar 2021.